

ILCSO's Manipulation of MARC Coding in Bib Records

This document is intended for the 44 former ILCSO libraries that once shared the LCS, FBR, and DRA systems. It describes some of the data manipulation that we did over the years, which explains some of the anomalies in data that these libraries may find in their systems today.

First, the document takes a chronological view of the data, showing where the data came from and how they were changed. Then we'll take a MARC view of the data as it appears today and explains some of the unexpected things that one might see.

LCS

LCS was a "Library Circulation System." We got the software from Ohio State University and modified it extensively. It had detailed records of holdings and circulation status, but its bibliographic data was confined to a few fields: main entry, title, publication date, publication place, and LC card number. Libraries joined LCS between 1978 and 1995. The state colleges and universities were among the first to join.

Some libraries had substantial files of MARC bib records when they joined. A few libraries migrated to LCS from other systems. But some libraries, especially the earlier and larger ones, did shelf list conversions to get their data into machine readable form. Sometimes there was authority control; sometimes not. Many shelf list records, of course, were pre-AACR2.

LCS data, whether they came from MARC files or not, were stored in all upper-case letters. The member libraries had separate files of bibliographic data. The unique identifier for the bibliographic record was the "title number". Title numbers were assigned sequentially as records entered each library's database. Each library was identified by a 2-letter campus code. For LCS records derived from MARC records the LCS title was only 245\$a.

Libraries handled analyzed series in a variety of ways in LCS. Some had a record for the series. Some had separate records for each piece. Some did both.

FBR

FBR stood for Full Bibliographic Records. We wanted a system to work with LCS that would allow for storage and searching of the rich fields in the catalog records. We licensed the software of a regional bibliographic utility, the Washington Library Network, so you will sometimes find reference to the system as "WLN". FBR stored full MARC records (upper and lower case), but had very little in the way of holdings data. So we added files to logically link FBR records to LCS records, and developed software to enable a searcher to move from an FBR record to the corresponding LCS holdings.

There was a team of catalog maintenance staff from ILCSO libraries who did data clean-up projects and authority control work on the FBR database. This benefited all of the ILCSO libraries. A set of data-sweeping tools was also developed which enabled us to make large numbers of corrections with a reasonable amount of labor.

Record Precedence in FBR

FBR was a union catalog with a single copy of each bib. Bib records were de-duplicated by OCLC control number. The most important thing to know about record precedence is that, in FBR, the version of a record on which your library's holdings appear might not be the version that your library produced. Indeed, it might have been produced by a public or special library. Generally speaking, the first copy of a record that reached FBR with a given OCLC number was the copy that was shared by all FBR users, with one exception. The ILC SO libraries could insert a special code in an OCLC record that would cause it to replace a record in FBR with the same control number.

It is not possible to reconstruct the sequence of loading records into FBR, but some general comments can be made. Initially, FBR could handle only the book, serial, and film formats. Later, it could handle scores and sound recordings. Maps, computer files and manuscripts came later. Initially, we loaded the records of UIU and the members of the River Bend Library System. Later we loaded the records of most OCLC users in Illinois, including many public and special libraries. Typically, when we decided to load a new category of records, we started by working them into the current weekly load, and then went back through our retrospective data starting from the earliest records. Special loads were done occasionally including UIUC CJK records from RLIN, Northwestern University Transportation library, University of Chicago, retrospective conversion projects, and OCLC major sets.

Tag Mapping in FBR

FBR had a million lines of code and sketchy documentation, so there was much that we never learned about it. This was particularly true of the indexing rules. So, especially at first, we took the approach of modifying the data, rather than the software, so that we would get the indexing features that we wanted.

In the leader, we mapped record type "h" (formerly used for microforms) to "a". We mapped bib levels "a" and "c" to "m". We mapped bib level "b" to "s". We mapped encoding levels "1" (one) and "i" (aye) to blank; we mapped encoding levels "e", "j", "k", "l" (el), "m", "w", and "7" to "1" (one). These mappings were made to accommodate FBR's MARC validation scheme.

The FBR software used the LC card number as its primary key for deduplicating. It used the OCLC number as a backup. We deleted LCCNs to get around this.

For a few years there were MARC 705 and 715 fields in music records for performing artists and ensembles. We mapped these to 700 and 710, and some years later OCLC did the same. The tagging was so similar that we're not aware that this mapping ever caused a problem.

We wanted to index the 028 (Publisher Number) field. We mapped it to 710 (Added Entry-Corporate Name) to accomplish this.

We may have wanted to index the 052 (Geographic classification) field. Our records are incomplete in this respect, but we may have mapped it to 651 (Subject Added Entry-Geographic Name).

We wanted to index the 246 (Varying Form of Title), 653 (Index Term-Uncontrolled), and 740 (Added Entry Uncontrolled Related/Analytical Title). We did this by mapping them to 247 (Former Title). For 653s, we did the mapping only for non-UK/MARC records and we changed the indicators to "02". Note that this gave us a hodge-podge of indicator values in the 247.

We wanted to index the 490 (Series Statement) with a first indicator of 0 (Series Not Traced). We mapped the tag to 440 (Series Statement) to accomplish this. Note that the 440 second indicator is a non-filing indicator, and we left this blank.

We wanted to index more subject headings. We mapped 652 (Subject Added Entry-Reversed Geographic) to 651, setting the second indicator to "7". We mapped 655 (Genre Heading) to 650.

We wanted to index the local subject headings in the 690-695 fields. We did this by mapping them to the corresponding 6xx fields thus:

690 -> 650

691 -> 651

692 -> 600

693 -> 610

694 -> 611

695 -> 630

The indicators and subfielding were reasonably compatible.

755 (Added Entry, Physical Characteristics) got mapped to 650.

Field Trimming in FBR

Long 040 (Cataloging Source) fields displayed poorly in FBR. We deleted subfields "b" and "e" and all occurrences of subfield "d" except the last.

Because FBR bib records were shared, we deleted all local notes (590 and 599 fields). Likewise, we deleted call numbers from 049, 089, 090, 091, 092, 096, and 099.

FBR could not display CJK characters, so we deleted 880 (Alternate Graphic Representation) fields from bibs.

FBR could not handle MARC records longer than 3050 bytes. Data entry staff were aware of this and often attempted to shorten records so that they could be loaded, but the loading programs still found records that were too long to be loaded. To deal with this, we first deleted the unwanted fields mentioned in the last few paragraphs. If the record was still too long, we deleted a set of fields. If it was still too long, we deleted another set of fields. If it was still too long, we deleted another set of fields. The fields deleted in each of these three passes were:

First pass:

- 009 Physical description field for archival collections
- 012 Terminal display
- 017 Copyright registration number
- 018 Copyright article-fee code
- 025 Overseas acquisitions number
- 039 Level of bibliographic control
- 044 Country of publishing code
- 051 LC copy, issue, offprint statement
- 055 Call numbers assigned in Canada
- 061 NLM copy statement
- 071 NAL copy statement
- 072 Subject category code
- 073 GPO item number

- 510 Citation note (deleted from serial records only)
- 512 Earlier/later volumes cataloged separately note
- 543 Solicitation information
- 583 Action code
- 584 Accumulation and frequency of use note
- 585 Exhibitions note
- 68x PRECIS codes
- 699 Added Class Number
- 765 Original language entry
- 767 Translation entry
- 851 Location
- 886 Foreign MARC information

Second pass:

- 242 Translation of title by cataloging agency
- 243 Collective uniform title
- 524 Preferred citation of described materials note
- 570 Editor note
- 760 Main series entry
- 762 Subseries entry
- 77x Misc.

Third pass:

- 033 Date, time and place of an event
- 043 Geographic area code
- 045 Time period of content
- 047 Form of musical composition code
- 048 Number of musical instruments or voices
- 257 Country of producing entity for archival films
- 265 Source for acquisition/subscription address
- 306 Playing time
- 350 Price
- 503 Bibliographic history note
- 505 Contents Note
- 518 Date, time and place of event note
- 520 Summary, etc. note
- 545 Bibliographical or historical note
- 547 Former title complexity note
- 555 Cumulative index/finding aids note
- 561 Provenance (deleted from AV records only)
- 567 Methodology note
- 581 Publications about described materials note
- 582 Related computer files note
- 787 Nonspecific relationship entry

"Phase 1"

We had a plan to restructure LCS and FBR, of which only Phase 1 was ever completed. Up till this point - 1990 - LCS was a pretty accurate representation of a library's holdings, but the searching left much to be desired. FBR had superior searching, but many LCS records were not represented there.

Phase 1 involved identifying the LCS records that did not have a corresponding FBR record, and making them into MARC records, which we called MARCettes. We compared these against FBR. If there was a full record in FBR to match, we would link it to LCS and discard the MARCette. If we could not find a match, we loaded the MARCette into FBR.

Although the rest of the restructuring plan was never completed, Phase 1 nonetheless brought us short-term improvements in searching and positioned us for our eventual move to DRA.

Note that, when we linked an FBR record to an LCS record, we never set the holdings in OCLC. Consequently, many libraries have holdings on OCLC MARC record in Voyager that are not represented in OCLC.

MARCettes, like LCS records, were in all upper case letters. However, before we loaded the MARCettes into FBR, we sent them to UTLAS for authority control. UTLAS was getting out of the authority control business and we were one of their last customers. Between their time constraints and ILCISO monetary constraints, not all MARCettes got authority control. But one still finds MARCettes upper case titles and publication places, but upper/lower case main entries.

When the MARCettes were loaded into FBR, we did not have OCLC numbers to use for deduplicating and we didn't want to deduplicate these records. So we created a control number from the 2-letter lower case LCS library code, followed by the LCS title number padded with leading zeroes. Since that time, libraries have done projects to upgrade most of the MARCettes to full MARC, but one will still find these control numbers in some records.

The creation of MARCettes was mostly pretty reasonable. The only surprise one might have in looking at a MARCette today is that we put that call number in a 699. The complete spec for Phase 1 is in the document "LCS Restructuring Plan Phase 1 Adding LCS Records to FBR."

Here are the record counts from Phase 1

Current library name	2-letter LCS campus code	LCS titles back then	LCS recs linked to full MARC bibs in FBR	MARCettes added to FBR
Aurora U	ar	85,669	1,082	924
Benedictine U	ib	87,734	3,569	2,444
Columbia C	cl	82,672	779	1,482
Chicago State U	cs	238,291	804	11,600
Catholic Theological Union	ct	60,119	104	98
Dominican U	ro	129,541	2,028	981
DePaul U	dp	275,966	1,328	3,714
Eastern IL U	ea	345,557	983	10,235
Elmhurst C	el	134,488	1,112	809
Governors State U	gs	216,363	2,527	2,981
IL Institute of Technology	it	181,368	4,348	4,753
IMSA	im	20,379	498	314

IL State U	is	714,638	171,958	199,153
Il Wesleyan U	iw	128,785	1,282	1,567
Judson C	ju	65,527	438	841
Kankakee Community C	kk	33,907	22,821	11,082
Lake Forest C	lf	165,026	582	3,160
Milliken U	ml	125,720	233	2,078
North Central C	nc	74,885	661	450
Northeastern IL U	ne	302,420	2,340	8,494
Northern IL U	ni	800,496	37,932	103,279
Oakton Community C	oa	69,456	677	624
Roosevelt U	ru	173,510	455	439
School of the Art Inst	sa	26,466	74	59
SIU-Carbondale	sc	1,050,250	16,287	97,818
SIU-Edwardsville	se	482,484	169,192	191,103
SIU-School of Medicine	sm	51,198	378	2,779
St. Xavier U	sx	86,359	1,252	1,371
Triton C	tc	68,235	3,399	1,928
UI-Chicago	cc	548,452	15,137	19,120
UI-Chicago	mc	131,142	24,833	52,438
UI-Springfield	ss	332,819	59,062	89,121
UI-Urbana	uc	3,395,480	316,511	1,678,714
Western IL U	we	446,546	1,564	2,314

DRA

In 1998, we migrated the data from LCS and FBR into DRA. We extracted the LCS holdings into a flat file and worked with DRA to make them into MFHDs and items. Bib records that had no corresponding holdings from LCS were abandoned. In most cases, these were bibs owned by ILLINET libraries, but not by ILC SO libraries.

At the time of the DRA conversion, UIC was running a NOTIS system as well as LCS and FBR. The bib records for UIC came from NOTIS. SIU School of Medicine was running the Georgetown LIS system in addition to LCS and FBR. Their bib records were extracted from that system.

Because bib records were shared by libraries in both FBR and DRA, we deleted all local notes (590 and 599) from the bib records. We'd been doing this for a while in FBR, but some of the older FBR records still had these fields.

Tag Mapping Reversal during the Conversion to DRA

Because DRA's indexing was more profile-able than FBR's, we made an effort to undo some of the tag mapping that we did in FBR. during the conversion to DRA.

We had been mapping 246, 653, and 740 into 247. To reverse this, if the second indicator was blank, we changed the tag to 653. If the bib level was not "s", we changed the tag to 740. Otherwise, we left the tag as 247. This reversal was far from perfect, so we decided to index these fields together, thus minimizing the chaos for library patrons. In particular, this processed retagged too many fields as 653s. Today this causes indexing problem in Voyager. See

“Correcting ‘the 653 problem’ how to search for it, recognize it from a printout, and correct it” at <http://www.carli.illinois.edu/mem-prod/I-Share/cat/corr653prob.html> .

We had been mapping 028 into 710. To reverse this, we examined the first 3 characters of \$a of 710s. If there were at least 2 digits there, we changed the tag to 028.

Field Trimming Reversal during the Conversion to DRA

Because DRA could handle longer MARC records than FBR, we made an effort to recover some of the data that had been trimmed out of the records. We did this by going back through the “ILLINET Tapes,” the files of OCLC records that we stored on behalf of most of the ILLINET libraries that used OCLC. There were several variations in how we handled the recovery of these data.

Take the 505 (Contents Note) for example. We went through the ILLINET tapes and created tiny MARC records consisting of the OCLC# and the 505(s). We compared records with the same OCLC# and selected the one with the longest 505(s). The 505s were then restored to the bib record that had the corresponding OCLC#. The following fields received the same treatment as the 505:

- 503 Bibliographic history note
- 518 Date, time and place of event note
- 520 Summary, etc. note
- 545 Bibliographical or historical note
- 547 Former title complexity note
- 555 Cumulative index/finding aids note
- 561 Provenance (was deleted from AV records only)
- 567 Methodology note
- 581 Publications about described materials note
- 582 Related computer files note
- 880 Alternate Graphic Representation

A variation on this was the treatment of the 760 (Main Series Entry) and 787 (Nonspecific Relationship Entry) fields. These fields were still present in some of the FBR records. We created tiny MARC records, as described in the previous paragraph, but we restored the fields to the bib records only when there were no fields with the same tag in the records.

We wanted to restore as many control numbers as possible, so we created similar tiny MARC records for the following fields, but we used them to restore each unique control number value back into the bib records.

- 010 LC Card Number
- 074 GPO Item Number

Record Precedence in DRA

Like FBR, DRA was a union catalog with a single bib record shared by all of the libraries that owned the title. The records that were copied from FBR to DRA were the ones that had holdings for ILC SO libraries on them. The other bibs were simply abandoned.

Like FBR, the first copy of a record that reached DRA with a given OCLC number was the copy that was shared by all FBR users, with one exception. The ILC SO libraries could insert a special

code in an OCLC record that would cause it to replace a record in FBR with the same control number.

Voyager

When we migrated from DRA to Voyager in 1992, we didn't do any special manipulation of bibliographic data. Whew! However, we did have to split up the bib records, sending the correct records to each database. There were some bib records that we could not assign to a database, so we put these in RESdb.

Field by Field

There are still some MARCettes, which were created in Phase 1. You'll know them by

- the title and place of publication are all capital letters
- the call# is in the 699 field

009

deleted from long FBR records during the first pass
never recovered

010

deleted from FBR records
restored during the conversion to DRA

012

deleted from long FBR records during the first pass
never recovered

017

deleted from long FBR records during the first pass
never recovered

018

deleted from long FBR records during the first pass
never recovered

025

deleted from long FBR records during the first pass
never recovered

028

mapped to 710 in FBR
mapped back during the conversion to DRA

033

deleted from long FBR records during the third pass
never recovered

039

deleted from long FBR records during the first pass
never recovered

040

In FBR, \$b, \$e, and all but the last \$d were trimmed

043

deleted from long FBR records during the third pass
never recovered

044

deleted from long FBR records during the first pass
never recovered

045

deleted from long FBR records during the third pass
never recovered

047

deleted from long FBR records during the third pass
never recovered

048

deleted from long FBR records during the third pass
never recovered

049

deleted from FBR records
never recovered

051

deleted from long FBR records during the first pass
never recovered

052

may have been mapped to 651 in FBR

055

deleted from long FBR records during the first pass
never recovered

061

deleted from long FBR records during the first pass
never recovered

071

deleted from long FBR records during the first pass
never recovered

072

deleted from long FBR records during the first pass
never recovered

073
deleted from long FBR records during the first pass
never recovered

074
deleted from FBR records
restored during the conversion to DRA

089
deleted from FBR records
never recovered

090
deleted from FBR records
never recovered

091
deleted from FBR records
never recovered

092
deleted from FBR records
never recovered

096
deleted from FBR records
never recovered

099
deleted from FBR records
never recovered

242
deleted from long FBR records during the second pass
never recovered

243
deleted from long FBR records during the second pass
never recovered

246
mapped to 247 in FBR
some were mapped back during the conversion to DRA

247
the 246, 653, and 740 tags were mapped to this in FBR
some were mapped back during the conversion to DRA

257
deleted from long FBR records during the third pass
never recovered

265

deleted from long FBR records during the third pass
never recovered

306

deleted from long FBR records during the third pass
never recovered

350

deleted from long FBR records during the third pass
never recovered

440

490 with Ind1=0 were mapped to 440 in FBR
no attempt was made to map these back during the conversion to DRA
a focus of data cleanup efforts

490

mapped to 440 in FBR
no attempt was made to map these back during the conversion to DRA
a focus of data cleanup efforts

503

deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

505

deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

510

deleted from long FBR records during the first pass
never recovered

512

deleted from long FBR records during the first pass
never recovered

518

deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

520

deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

524

deleted from long FBR records during the second pass
never recovered

543

deleted from long FBR records during the first pass
never recovered

545
deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

547
deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

555
deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

561
deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

567
deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

570
deleted from long FBR records during the second pass
never recovered

581
deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

582
deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

583
deleted from long FBR records during the first pass
never recovered

584
deleted from long FBR records during the first pass
never recovered

585
deleted from long FBR records during the first pass
never recovered

590
deleted from most FBR records
deleted from the remaining records during the conversion to DRA

599

deleted from most FBR records
deleted from the remaining records during the conversion to DRA

600
692 fields were mapped to 600 in FBR

610
693 fields were mapped to 610 in FBR

611
694 fields were mapped to 611 in FBR

630
695 fields were mapped to 630 in FBR

650
655 fields were mapped to 650 in FBR
690 fields were mapped to 650 in FBR
755 fields were mapped to 650 in FBR

651
052 fields may have been mapped to 651 in FBR
652 fields were mapped to 651 in FBR
691 fields were mapped to 651 in FBR

652
mapped to 651 fields and Ind2 set to "7" in FBR

653
mapped to 247 in FBR
some were mapped back during the conversion to DRA

655
mapped to 650 in FBR

68x
deleted from long FBR records during the first pass
never recovered

690
mapped to 650 in FBR

691
mapped to 651 in FBR

692
mapped to 600 in FBR

693
mapped to 610 in FBR

694

mapped to 611 in FBR

695

mapped to 630 in FBR

699

deleted from most FBR records

700

705 fields were mapped to 700 in FBR

705

mapped to 700 in FBR

710

715 fields were mapped to 710 in FBR

028 fields were mapped to 710 in FBR

some 028s were mapped back during the conversion to DRA

715

mapped to 710 in FBR

740

mapped to 247 in FBR

some were mapped back during the conversion to DRA

755

mapped to 650 in FBR

760

deleted from long FBR records during the second pass
recovered, in part, during the conversion to DRA

762

deleted from long FBR records during the second pass
never recovered

765

deleted from long FBR records during the first pass
never recovered

767

deleted from long FBR records during the first pass
never recovered

77x

deleted from long FBR records during the second pass
never recovered

787

deleted from long FBR records during the third pass
recovered, in part, during the conversion to DRA

851
deleted from long FBR records during the first pass
never recovered

880
deleted from all FBR records
recovered, in part, during the conversion to DRA

886
deleted from long FBR records during the first pass
never recovered